Package 'msaeDB'

December 31, 2020

Type Package

Title Difference Benchmark for Multivariate Small Area Estimation

Version 0.1.0

Author Zaza Yuda Perwira, Azka Ubaidillah

Maintainer Zaza Yuda Perwira <221710086@stis.ac.id>

Description This Package is implementing Benchmarking Method for Multivariate Small Area Estimation under Fay Herriot Model.
Multivariate Small Area Estimation (MSAE) is a development of Univariat Small Area Estimation that
considering the correlation among response variables and borrowing the strength from auxiliary variables
effectiveness of a domain sample size, the multivariat model in this package is based on Multivariate
model 1 proposed by Roberto Benavent and Domigo Morales (2015) <DOI: 10.1016/j.csda.2015.07.013.>.
Benchmarking in Small Area Estimation is a modification of Small Area Estimation model to guarantees that the
aggreagate weighted mean of the county predictors equals the corresponding weighted mean of survey estimates.
Difference Benchmarking is the simplest but widely used by multiplying EBLUP estimator by the common adjustment
factor (J.N.K Rao and Isabela Molina, 2013).

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

URL https://github.com/zazaperwira/msaeDB

BugReports https://github.com/zazaperwira/msaeDB/issues

Suggests knitr, rmarkdown, covr

VignetteBuilder knitr

Imports MASS, magic, stats

Depends R (>= 2.10)

R topics documented:

`datamsaeDB` *Sample Data for Multivariat Small Area Estimation with Difference Benchmarking*

Description

Dataset to simulate Difference benchmarking of Multivariate Fay Herriot model

This data is generated base on multivariate Fay Herriot Model by these following steps:

1. Generate explanatory variables `X1` and `X2`. Take μ_{X1} = μ_{X1} = 10, σ_{X11}=1, σ_{X2}=2, and ρ_x= 1/2.
   Sampling error e is generated with the following σ_{e11} = 0.15, σ_{e22} = 0.25,σ_{e33} = 0.35, and ρ_e = 1/2.
   For random effect u, we set σ_{u11}= 0.2, σ_{u22}= 0.6, and σ_{u33}= 1.8.
   for the weight we generate `w1` `w2` `w3` by set the `w1 ~ U(25,30)` , `w2 ~ U(25,30)`, `w3 ~ U(25,30)`
   Calculate direct estimation `Y1` `Y2` `Y3` where $Y_i = X * \beta + u_i + e_i$
2. Then combine the direct estimations `Y1` `Y2` `Y3`, explanatory variables `X1` `X2`, weights `w1` `w2` `w3`, and sampling varians covarians `v1` `v12` `v13` `v2` `v23` `v3` in a dataframe then named as datamsaeDB

Usage

```
datamsaeDB
```

Format

A data frame with 30 rows and 14 variables:

Y1 Direct Estimation of Y1

Y2 Direct Estimation of Y2

Y3 Direct Estimation of Y3

X1 Auxiliary variable of X1

X2 Auxiliary variable of X2

w1 sampling weight of Y1

w2 sampling weight of Y2

w3 sampling weight of Y3

v1 Sampling Variance of Y1

v12 Sampling Covariance of Y1 and Y2

v13 Sampling Covariance of Y1 and Y3

v2 Sampling Variance of Y2

v23 Sampling Covariance of Y2 and Y3

v3 Sampling Variance of Y3

`msaedb` *EBLUPs under Multivariate Fay Herriot Model with Difference Benchmarking*

Description

This function produces EBLUPs, MSE, and Aggregarion of Multivariat SAE with Difference Benchmarking

Usage

```
msaedb(
  formula,
  vardir,
  weight,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

`formula`	list of formula that describe the fitted model
`vardir`	Sampling variance of direct estimation, if the data is included in data frame so it is the vector with the name of sampling variances. if it is not, it is a data frame of sampling variance in order : `var1,cov12,.,cov1r,var2,cov23,.,cov2r.,.cov(r-1)(r),var(r)`
`weight`	Vector of proportion of units in small area
`samevar`	Whether the variances of the data is same or not. Logical input with default `FALSE`
`MAXITER`	Maximum number of iteration in Fisher-scoring algorithm with default `100`
`PRECISION`	Limit of fisher-scoring convergence tolerance with default `1e-4`
`data`	The data frame

Value

This function returns a list of the following objects:

`MSAE_Eblup`	a dataframe with the values of the EBLUPs estimators
`MSE_Eblup`	a dataframe with the values of estimated mean square errors of EBLUPs estimators
`fit`	a list containing the following objects:

- method : The fitting method (this function is using "REML")
- convergence : the convergence result of fisher scoring algorithm (Logical Value)

- iterations : The number of Fisher-Scoring algorithm iterations
- estcoef : a dataframe with the estimated model coefficient, standard error,t statistics, p-values of the significance of each coefficient
- refvar : a dataframe with estimated random effect variance
- informationFisher : a matrix of information fisher from Fisher-scoring algorithm

`difference_benchmarking`
a list containing the following objects:

- Estimation : a dataframe with the value of Benchmarked EBLUPs estimators
- Aggregation : The Aggregation of benchmarked EBLUPs estimations, EBLUPs Estimations and direct estimations
- MSE_DB : a dataframe with the values of estimated mean square errors of benchmarked EBLUPs estimators
- g.4a : first component of g4 in difference benchmarking MSE estimation formula
- g.4b : second component of g4 in difference benchmarking MSE estimation formula

Examples

```
##load dataset
data(datamsaeDB)

#Compute Fitted model for Y1, Y2, and Y3
#Y1 ~ X1 + X2
#Y2 ~ X2
#Y3 ~ X1

##Using parameter 'data'
formula = list(f1 = Y1~X1+X2,
               f2 = Y2~X2,
               f3 = Y3~X1)
vardir = c("v1","v12","v13","v2","v23","v3")
weight = c("w1","w2","w3")
msaeDB <- msaedb(formula, vardir, weight, data=datamsaeDB)

##Do not use parameter 'data'
formula = list(f1 = datamsaeDB$Y1~datamsaeDB$X1+datamsaeDB$X2,
               f2 = datamsaeDB$Y2~datamsaeDB$X2,
               f3 = datamsaeDB$Y3~datamsaeDB$X1)
vardir = datamsaeDB[,c("v1","v12","v13","v2","v23","v3")]
weight = datamsaeDB[,c("w1","w2","w3")]
msaeDB <- msaedb(formula, vardir, weight)

msaeDB$MSAE_Eblup       #to see EBLUP Estimators
msaeDB$MSE_Eblup        #to see estimated MSE of EBLUP estimators
msaeDB$difference_benchmarking$Estimation   #to see Benchmarked EBLUP Estimators
msaeDB$difference_benchmarking$MSE_DB     #to see estimated MSE of Benchmarked EBLUP Estimators
msaeDB$difference_benchmarking$Aggregation  #to see the aggregation of, benchmarking.
```

`msaefh` *EBLUPs under Multivariate Fay Herriot Model*

Description

This function produces EBLUPs, MSE of Multivariat SAE

Usage

```
msaefh(
  formula,
  vardir,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

`formula`	list of formula that describe the fitted model
`vardir`	Sampling variance of direct estimation, if the data is included in data frame so it is the vector with the name of sampling variances. if it is not, it is a data frame of sampling variance in order : `var1,cov12,.,cov1r,var2,cov23,.,cov2r.,.cov(r-1)(r),var(r)`
`samevar`	Whether the variances of the data is same or not. Logical input with default `FALSE`
`MAXITER`	Maximum number of iteration in Fisher-scoring algorithm with default `100`
`PRECISION`	Limit of fisher-scoring convergence tolerance with default `1e-4`
`data`	The data frame

Value

This function returns a list of the following objects:

`MSAE_Eblup`	a dataframe with the values of the EBLUPs estimators
`MSE_Eblup`	a dataframe with the values of estimated mean square errors of EBLUPs estimators
`fit`	a list containing the following objects:

- method : The fitting method (this function is using "REML")
- convergence : the convergence result of fisher scoring algorithm (Logical Value)
- iterations : The number of Fisher-Scoring algorithm iterations
- estcoef : a dataframe with the estimated model coefficient, standard error,t statistics, p-values of the significance of each coefficient
- refvar : a dataframe with estimated random effect variance
- informationFisher : a matrix of information fisher from Fisher-scoring algorithm

Examples

```
##load dataset
data(datamsaeDB)

#Compute Fitted model for Y1, Y2, and Y3
#Y1 ~ X1 + X2
#Y2 ~ X2
#Y3 ~ X1

##Using parameter 'data'
formula = list(f1 = Y1~X1+X2,
               f2 = Y2~X2,
               f3 = Y3~X1)
vardir = c("v1","v12","v13","v2","v23","v3")
msaeFH <- msaefh(formula, vardir, data=datamsaeDB)

#Do not use parameter 'data'
formula = list(f1 = datamsaeDB$Y1~datamsaeDB$X1+datamsaeDB$X2,
               f2 = datamsaeDB$Y2~datamsaeDB$X2,
               f3 = datamsaeDB$Y3~datamsaeDB$X1)
vardir = datamsaeDB[,c("v1","v12","v13","v2","v23","v3")]
msaeFH <- msaefh(formula, vardir)

msaeFH$MSAE_Eblup       #to see EBLUP Estimators
msaeFH$MSE_Eblup        #to see estimated MSE of EBLUP estimators
```

---

usaedb — *EBLUPs under Univariate Fay Herriot Model with Difference Benchmarking*

Description

This function produces EBLUPs, MSE, and Aggregarion of Univariat SAE with Difference Benchmarking

Usage

```
usaedb(
  formula,
  vardir,
  weight,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

formula	list of formula that describe the fitted model

vardir	Sampling variance of direct estimation, if the data is included in data frame so it is the vector with the name of sampling variances. if it is not, it is a data frame of sampling variance in order : var1,cov12,.,cov1r,var2,cov23,.,cov2r.,.cov(r-1)(r),var(r)
weight	Vector of proportion of units in small area
samevar	Whether the variances of the data is same or not. Logical input with default FALSE
MAXITER	Maximum number of iteration in Fisher-scoring algorithm with default 100
PRECISION	Limit of fisher-scoring convergence tolerance with default 1e-4
data	The data frame

Value

This function returns a list of the following objects:

USAE_Eblup	a dataframe with the values of the EBLUPs estimators
MSE_Eblup	a dataframe with the values of estimated mean square errors of EBLUPs estimators
fit	a list containing the following objects:

- method : The fitting method (this function is using "REML")
- convergence : the convergence result of fisher scoring algorithm (Logical Value)
- iterations : The number of Fisher-Scoring algorithm iterations
- estcoef : a dataframe with the estimated model coefficient, standard error,t statistics, p-values of the significance of each coefficient
- refvar : a dataframe with estimated random effect variance
- informationFisher : a matrix of information fisher from Fisher-scoring algorithm

difference_benchmarking
: a list containing the following objects:

- Estimation : a dataframe with the value of Benchmarked EBLUPs estimators
- Aggregation : The Aggregation of benchmarked EBLUPs estimations, EBLUPs Estimations and direct estimations
- MSE_DB : a dataframe with the values of estimated mean square errors of benchmarked EBLUPs estimators
- g.4a : first component of g4 in difference benchmarking MSE estimation formula
- g.4b : second component of g4 in difference benchmarking MSE estimation formula

Examples

```
##load dataset
data(datamsaeDB)

#Compute Fitted model for Y1, Y2, and Y3
#Y1 ~ X1 + X2
#Y2 ~ X2
#Y3 ~ X1

##Using parameter 'data'
formula = list(f1 = Y1~X1+X2,
```

```
               f2 = Y2~X2,
               f3 = Y3~X1)
vardir = c("v1","v12","v13","v2","v23","v3")
#Note : in real data for univariate SAE, if you does not have the valuse of covariances,
#       set covariancse as zero in the dataframe

weight = c("w1","w2","w3")
usaeDB <- usaedb(formula, vardir, weight, data=datamsaeDB)

##Do not use parameter 'data'
formula = list(f1 = datamsaeDB$Y1~datamsaeDB$X1+datamsaeDB$X2,
               f2 = datamsaeDB$Y2~datamsaeDB$X2,
               f3 = datamsaeDB$Y3~datamsaeDB$X1)
vardir = datamsaeDB[,c("v1","v12","v13","v2","v23","v3")]
#Note : in real data for univariate SAE, if you does not have the valuse of covariances,
#       set covariancse as zero in the dataframe
weight = datamsaeDB[,c("w1","w2","w3")]
usaeDB <- usaedb(formula, vardir, weight = weight)

usaeDB$USAE_Eblup       #to see EBLUP Estimators
usaeDB$MSE_Eblup        #to see estimated MSE of EBLUP estimators
usaeDB$difference_benchmarking$Estimation   #to see Benchmarked EBLUP Estimators
usaeDB$difference_benchmarking$MSE_DB     #to see estimated MSE of Benchmarked EBLUP Estimators
usaeDB$difference_benchmarking$Aggregation  #to see the aggregation of, benchmarking
```

---

usaefh *EBLUPs under Univariate Fay Herriot Model*

Description

This function produces EBLUPs, MSE of Univariate SAE

Usage

```
usaefh(
  formula,
  vardir,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

formula	list of formula that describe the fitted model
vardir	Sampling variance of direct estimation, if the data is included in data frame so it is the vector with the name of sampling variances. if it is not, it is a data frame of sampling variance in order : var1,cov12,.,cov1r,var2,cov23,.,cov2r.,.cov(r-1)(r),var(r)
samevar	Whether the variances of the data is same or not. Logical input with default FALSE

MAXITER	Maximum number of iteration in Fisher-scoring algorithm with default 100
PRECISION	Limit of fisher-scoring convergence tolerance with default 1e-4
data	The data frame

Value

This function returns a list of the following objects:

USAE_Eblup	a dataframe with the values of the EBLUPs estimators
MSE_Eblup	a dataframe with the values of estimated mean square errors of EBLUPs estimators
fit	a list containing the following objects:

- method : The fitting method (this function is using "REML")
- convergence : the convergence result of fisher scoring algorithm (Logical Value)
- iterations : The number of Fisher-Scoring algorithm iterations
- estcoef : a dataframe with the estimated model coefficient, standard error,t statistics, p-values of the significance of each coefficient
- refvar : a dataframe with estimated random effect variance
- informationFisher : a matrix of information fisher from Fisher-scoring algorithm

Examples

```
##load dataset
data(datamsaeDB)

#Compute Fitted model for Y1, Y2, and Y3
#Y1 ~ X1 + X2
#Y2 ~ X2
#Y3 ~ X1

##Using parameter 'data'
formula = list(f1 = Y1~X1+X2,
               f2 = Y2~X2,
               f3 = Y3~X1)
vardir = c("v1","v12","v13","v2","v23","v3")
#Note : in real data for univariate SAE, if you does not have the valuse of covariances,
#       set covariancse as zero in the dataframe

usaeFH <- usaefh(formula, vardir, data=datamsaeDB)

##Do not use parameter 'data'
formula = list(f1 = datamsaeDB$Y1~datamsaeDB$X1+datamsaeDB$X2,
               f2 = datamsaeDB$Y2~datamsaeDB$X2,
               f3 = datamsaeDB$Y3~datamsaeDB$X1)
vardir = datamsaeDB[,c("v1","v12","v13","v2","v23","v3")]
#Note : in real data for univariate SAE, if you does not have the valuse of covariances,
#       set covariancse as zero in the dataframe
usaeFH <- usaefh(formula, vardir)

usaeFH$USAE_Eblup       #to see EBLUP Estimators
usaeFH$MSE_Eblup        #to see estimated MSE of EBLUP estimators
```

Index